



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2020

---

## **Domestication of high-copy transposons underlays the wheat small RNA response to an obligate pathogen**

Poretti, Manuel ; Praz, Coraline Rosalie ; Meile, Lukas ; Kälin, Carol ; Schaefer, Luisa Katharina ; Schläfli, Michael ; Widrig, Victoria ; Sanchez-Vallet, Andrea ; Wicker, Thomas ; Bourras, Salim

**Abstract:** Plant genomes have evolved several evolutionary mechanisms to tolerate and make use of transposable elements (TEs). Of these, transposon domestication into cis-regulatory and microRNA (miRNA) sequences is proposed to contribute to abiotic/biotic stress adaptation in plants. The wheat genome is derived at 85% from TEs, and contains thousands of miniature inverted-repeat transposable elements (MITEs), whose sequences are particularly prone for domestication into miRNA precursors. In this study, we investigate the contribution of TEs to the wheat small RNA immune response to the lineage-specific, obligate powdery mildew pathogen. We show that MITEs of the Mariner superfamily contribute the largest diversity of miRNAs to the wheat immune response. In particular, MITE precursors of miRNAs are wide-spread over the wheat genome, and highly conserved copies are found in the Lr34 and QPm.tut-4A mildew resistance loci. Our work suggests that transposon domestication is an important evolutionary force driving miRNA functional innovation in wheat immunity.

DOI: <https://doi.org/10.1093/molbev/msz272>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-184038>

Journal Article

Published Version




The following work is licensed under a Creative Commons: Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

Originally published at:

Poretti, Manuel; Praz, Coraline Rosalie; Meile, Lukas; Kälin, Carol; Schaefer, Luisa Katharina; Schläfli, Michael; Widrig, Victoria; Sanchez-Vallet, Andrea; Wicker, Thomas; Bourras, Salim (2020). Domestication of high-copy transposons underlays the wheat small RNA response to an obligate pathogen. *Molecular Biology and Evolution*, 37(3):839-848.

DOI: <https://doi.org/10.1093/molbev/msz272>

# Domestication of High-Copy Transposons Underlays the Wheat Small RNA Response to an Obligate Pathogen

Manuel Poretti <sup>1</sup>, Coraline Rosalie Praz,<sup>1</sup> Lukas Meile,<sup>2</sup> Carol Kälin,<sup>1</sup> Luisa Katharina Schaefer,<sup>1</sup> Michael Schläfli,<sup>1</sup> Victoria Widrig,<sup>1</sup> Andrea Sanchez-Vallet,<sup>2</sup> Thomas Wicker,<sup>\*,1</sup> and Salim Bourras<sup>\*,1,3</sup>

<sup>1</sup>Department of Plant and Microbial Biology, University of Zurich, Zurich, Switzerland

<sup>2</sup>Plant Pathology, Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

<sup>3</sup>Department of Forest Mycology and Plant Pathology, Division of Plant Pathology, Swedish University of Agricultural Sciences, Uppsala, Sweden

\*Corresponding authors: E-mails: wicker@botinst.uzh.ch; salim.bourras@slu.se.

Associate editor: Brandon Gaut

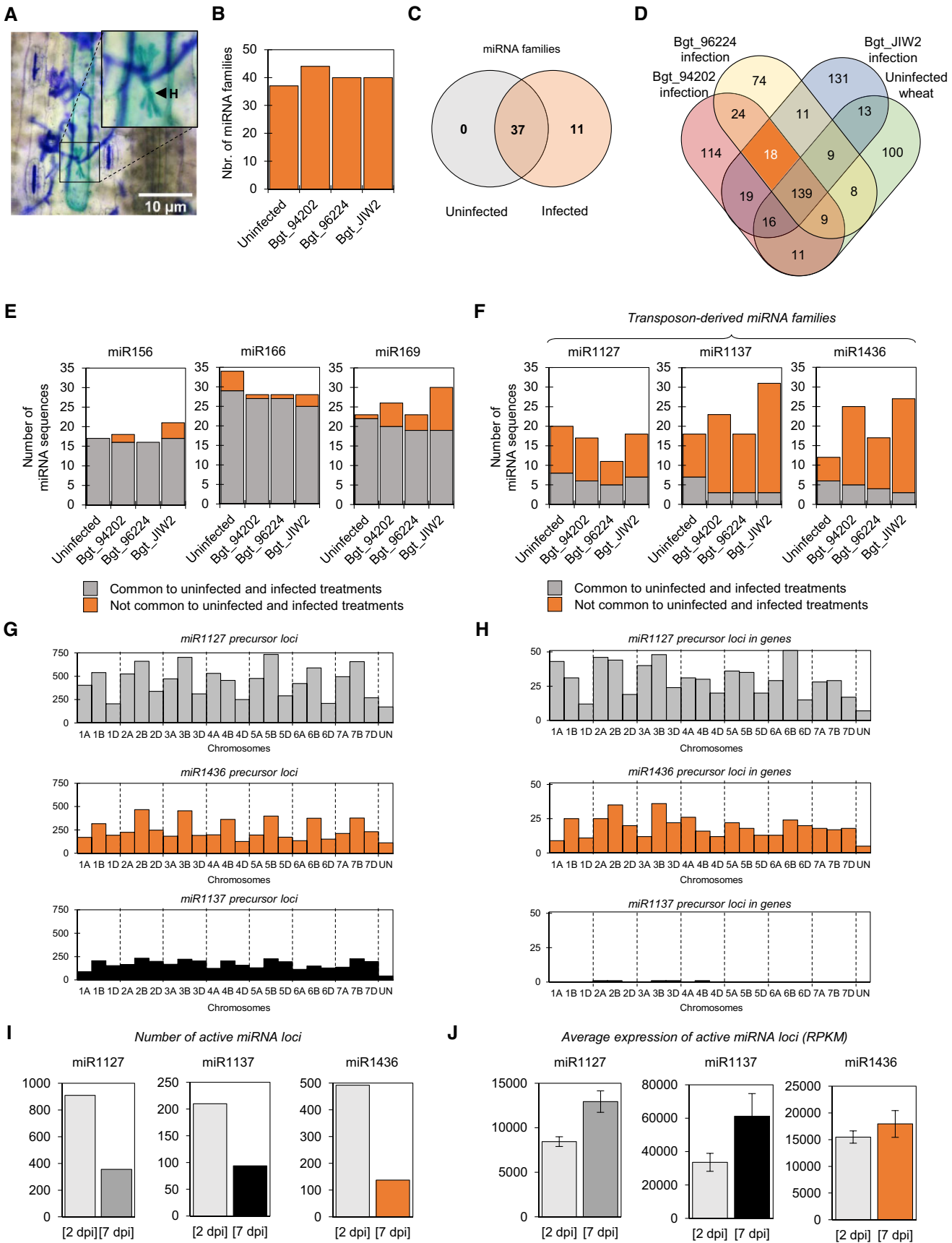
## Abstract

Plant genomes have evolved several evolutionary mechanisms to tolerate and make use of transposable elements (TEs). Of these, transposon domestication into cis-regulatory and microRNA (miRNA) sequences is proposed to contribute to abiotic/biotic stress adaptation in plants. The wheat genome is derived at 85% from TEs, and contains thousands of miniature inverted-repeat transposable elements (MITEs), whose sequences are particularly prone for domestication into miRNA precursors. In this study, we investigate the contribution of TEs to the wheat small RNA immune response to the lineage-specific, obligate powdery mildew pathogen. We show that MITEs of the *Mariner* superfamily contribute the largest diversity of miRNAs to the wheat immune response. In particular, MITE precursors of miRNAs are wide-spread over the wheat genome, and highly conserved copies are found in the Lr34 and QPm.tut-4A mildew resistance loci. Our work suggests that transposon domestication is an important evolutionary force driving miRNA functional innovation in wheat immunity.

**Key words:** wheat, transposable elements, small RNAs, powdery mildew.

Pathogens among other stresses, apply strong selective pressure on host genomes (Cagliani and Sironi 2013; Karasov et al. 2014). Transposable elements (TEs) are highly diversified mobile genetic units that contribute in a myriad of ways to genetic variation, genome evolution, and stress-adaptation (Rebollo et al. 2012; Vicent and Casacuberta 2017). In particular, TEs donate lineage-specific regulatory sequences, of which some can be domesticated into microRNA (miRNA) precursors (Mariño-Ramírez et al. 2005; Piriyaongsa et al. 2007; Roberts et al. 2014). Miniature inverted-repeat transposable elements (MITEs) of the “Stowaway” type (*Mariner* superfamily) (Bureau and Wessler 1994) were described as particularly good candidates for such domestication, as they terminate in a highly conserved 5'-CTCCCTCC motif which is repeated in reverse orientation at the 3' end (GGAGGGAG) (Bureau and Wessler 1994; Piriyaongsa and Jordan 2008; Yan et al. 2011). In cereals, RNA interference (RNAi) plays an important role in the regulation of stress-related responses including immunity (Budak et al. 2015). The 17 Gb genome of bread wheat (*Triticum aestivum* L.) is derived at 85% from TEs and contains over 100,000 *Stowaway* MITEs (Wicker et al. 2018). Thus, wheat is a particularly relevant system to study the contribution of TE domestication to adaptation and co-evolution with pathogens. In this study we address the contribution of TE derived miRNAs to the wheat response to the powdery mildew fungus *Blumeria graminis* f.sp. *tritici* (*B.g. tritici*), a highly specialized, lineage-specific, obligate parasite.

To identify the wheat miRNA response to powdery mildew, we generated four small RNA libraries from the susceptible wheat cultivar “Chinese Spring” corresponding to three “infected” treatments with three virulent *B.g. tritici* isolates Bgt\_96224, Bgt\_94202, and Bgt\_JIW2, and one “uninfected” control. Samples were harvested 2 days post inoculation (dpi), a stage corresponding to the formation of highly specialized feeding structures called “haustoria” (fig. 1A), which is a hallmark for successful host invasion (supplementary note S1, Supplementary Material online). The obtained deep small RNA sequencing data were applied to careful and stringent miRNA prediction using both homology-based and de novo annotation approaches (supplementary fig. S1, Supplementary Material online, supplementary note S1, Supplementary Material online), resulting in the identification of 696 unique miRNA sequences (i.e., sequence variants). Of these, 255 corresponded to potentially novel miRNAs, whereas 441 could be categorized into 48 previously known families, of which 16 are derived from TEs (fig. 1B, supplementary file S1, Supplementary Material online, supplementary table S1, Supplementary Material online). In total, 37 of the 48 families were commonly found between uninfected wheat and at least one of the infected samples. Interestingly, no miRNA family was found only present in uninfected plants, whereas 11 families were found only in the infected samples (fig. 1C, supplementary table S2, Supplementary Material online). This indicates that entire groups of miRNAs are



**FIG. 1.** Wheat miRNA identification from mildew infected and uninfected wheat samples. (A) Micrographs of a wheat powdery mildew “haustorium” feeding structure, formed on a single epidermal cell of a susceptible host at an early stage of infection. Light blue cell staining corresponds to transient expression of a GUS reporter serving as a contrastant that reveals the haustorium invagination. Epiphytic hyphae are

specifically expressed upon mildew infection. The findings are similar at the level of individual sequence variants. We found that circa 43% of the 696 identified miRNA sequence variants are encoded in the B subgenome ([supplementary table S3, Supplementary Material online](#)). In total, 205 were found in all four samples, 100 were specific to the uninfected control, whereas 391 were found in at least one sample of mildew-infected wheat but absent in the uninfected control ([fig. 1D](#)). We have also assessed if the wheat miRNA response to mildew is based on specific induction of different family members in infected and uninfected samples. We therefore selected six miRNA families for which at least 10 isoforms were predicted (miR156, miR166, miR169, miR1127, miR1137, miR1436) and compared the presence of individual family members in infected versus uninfected wheat ([fig. 1E, F, supplementary table S4, Supplementary Material online](#)). We found that the miR156, miR166, and miR169 variants induced upon infection were very similar and strongly overlapping with those already present in the uninfected wheat ([fig. 1E](#)). In contrast, we found that distinctly different variants of the miR1127, miR1137, and miR1436 families are induced upon infection and only a few were also found in the uninfected treatment ([fig. 1F, supplementary table S4, Supplementary Material online](#)). A closer look to the origin of the six families revealed that miR1127, miR1137, and miR1436 are all encoded within TEs, but not miR156, miR166, and miR169. Moreover, all these TE encoded families are derived from the “Mariner” superfamily MITEs of the *Stowaway* type ([supplementary table S1, Supplementary Material online](#)). Interestingly, these three TE-derived miRNA families were previously suggested to be involved in abiotic and biotic stress responses ([Djami-Tchatchou et al. 2017; Ravichandran et al. 2019](#)). The three families that showed only little variation between infected and uninfected samples (miR156, miR166, and miR169) are well-known regulators of development in plants ([Jung and Park 2007; Wu et al. 2009; Sorin et al. 2014; Xu et al. 2014](#)). They are found in all samples as highly abundant sequences ([supplementary table S5, Supplementary Material online](#)), indicating that they are expressed in the entire leaf irrespective of mildew presence. In contrast, the three transposon-derived miRNA families (miR1127, miR1137, and miR1436), are represented only by a small number of reads in the raw sequence data ([supplementary table S5, Supplementary Material online](#)), probably reflecting the fact that they are only accumulating in the small number of epidermal cells infected by mildew at 2 dpi ([supplementary](#)

[note S1, Supplementary Material online](#)). Notably, some miRNA sequence variants were found in all three infected samples but were absent in uninfected samples (examples in [supplementary table S6, Supplementary Material online](#)).

Since TEs are the most rapidly evolving of all genomic sequences, it was suggested that transposon domestication may provide a mechanism for the emergence of lineage-specific miRNA genes. We therefore assessed the phylogenetic conservation of the precursor sequences of the same six miRNA families in other plant species. We found that miR156, miR166, and miR169 are encoded by phylogenetically conserved precursor sequences that can be found in barley, *Brachypodium*, rice, and maize ([supplementary fig. S2, Supplementary Material online](#)). However, miR1127, miR1137, and miR1436 were specifically encoded by sequences only found in the “*tritici*” tribe that were absent in other species ([supplementary fig. S2B, Supplementary Material online, supplementary files S2–S4, Supplementary Material online](#)). These are unevenly distributed over the wheat subgenomes, and those encoding miR1137 are quasi exclusively encoded outside of genes ([fig. 1G, H, supplementary note S2, Supplementary Material online, supplementary fig. S3, Supplementary Material online, supplementary files S5–S7, Supplementary Material online](#)). We therefore conclude that the wheat miRNA response to mildew is based on the induction of specific members from genetically diverse miRNA families, encoded within lineage specific transposons with contrasting genome-wide distribution patterns.

In order to further investigate the activity of these miRNA families, we compared expression levels of their precursors in RNA-Seq data from the wheat cultivar “Chinese Spring” infected with isolate Bgt\_96224 (i.e., the same genetic material used for miRNA annotation) at 2 dpi ([Praz et al. 2018](#)) and at 7 dpi (see Materials and Methods, [supplementary note S2, Supplementary Material online](#)). We found that the total number of active miRNA loci (i.e., showing expression, [supplementary note S2, Supplementary Material online](#)) decreased at later (7 dpi) as compared with early (2 dpi) infection stages ([fig. 1I](#)). However, the loci active at 7 dpi showed a higher average expression than those active at 2 dpi ([fig. 1J](#)). This suggests that disease progression is probably imposing a tradeoff between the total number of miRNA loci that are activated and the maximum level of expression that is reached per locus.

We have also found that TE-derived miRNAs had most often homology to high-copy *Stowaway* MITE elements of

**FIG. 1.** Continued

stained with coomassie blue. (B, C) Comparison of miRNA families predicted from mildew infected and noninfected wheat samples. (B) Number of miRNA families predicted from each sample. (C) Venn diagram comparing the miRNA families found in infected versus uninfected samples. (D) Venn diagram comparing miRNA sequences induced in all treatments. A subset of 18 miRNAs found in all infected samples but never in the control is highlighted in orange. (E) Stacked bar plot illustrating strong overlap between the miR156, miR166, and miR169 sequence variants induced upon mildew infection versus the control. (F) Stacked bar plot showing only partial to little overlap between the miR1127, miR1137, miR1436 sequence variants present in infected versus uninfected wheat. (G) Genome wide distribution of the miR1127, miR1137, miR1436 precursors. (H) Genome wide distribution of the miR1127, miR1137, miR1436 precursors encoded within genes. (I) Comparison of the number of active miR1127, miR1137, miR1436 precursor loci at early (2 dpi) versus advanced (7 dpi) infection stages. (J) Average expression levels of the subset of active miRNA loci described in (I). Error bars correspond to the standard error of the mean. See [supplementary note S2, Supplementary Material online](#).

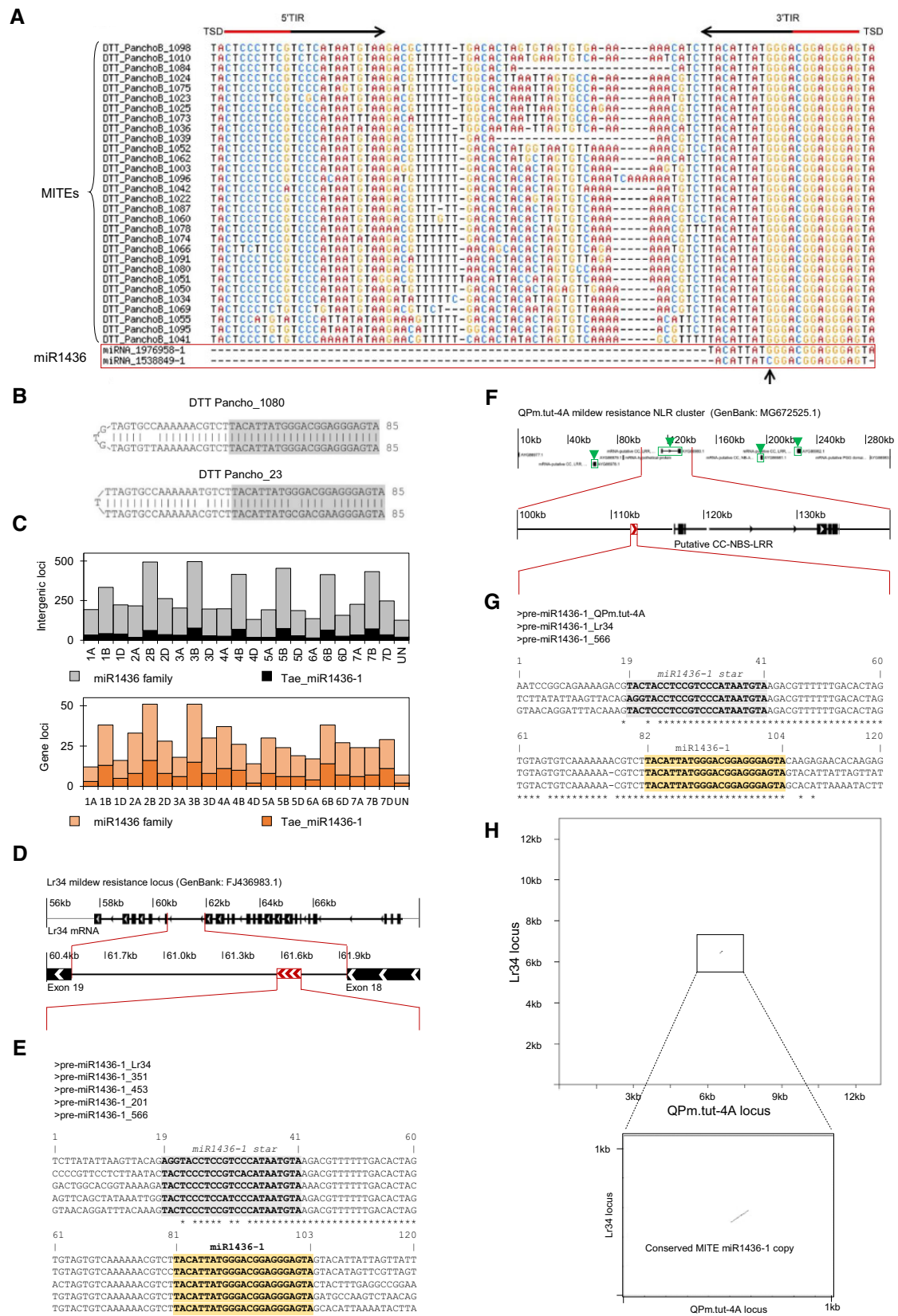


the “*Mariner*” superfamily of DNA transposons (supplementary table S2, Supplementary Material online), and they are frequently found inside their terminal inverted repeats (TIRs, fig. 2A, supplementary fig. S4, Supplementary Material online, supplementary note S3, Supplementary Material online). Furthermore, 103 of 131 (79%) are specifically induced upon infection (supplementary fig. S4, Supplementary Material online). For these sequences, highly stable RNA secondary structures can be predicted where miRNA variants form the ends of a stem loop spanning the miRNA and the complementary miRNA star sequence (fig. 2B). Because of sequence diversity within the MITE family as well as differences between the 5′ and 3′ TIRs, stem loops with mismatches between the predicted miRNA and star sequences can occur (fig. 2B). In the case of miR1436, ~70% of the *DTT\_Pancho* copies examined can form perfect hairpins without any mismatches in the region corresponding to the miR1436 sequence, whereas ~30% have at least 1 mismatch between the 5′ and the reverse complement of the 3′ TIR (fig. 2B). The finding that the largest diversity of miRNA sequences is contributed by TEs, and that many of these are encoded by MITEs, suggest these might play important roles in host–pathogen co-evolution through the regulation of the wheat immune response. We therefore focused on deeper characterization of MITE encoded miRNAs from the subset of 18 miRNA sequences commonly found in all three infected samples but not in the uninfected wheat control (fig. 1D). Of these, four are derived from MITEs and have the highest number of predicted miRNA precursor loci in the genome (supplementary table S7, Supplementary Material online). Two sequence variants were derived from *DTT\_Pancho* MITEs family miR1436. Although sequence variant Tae\_miR1436-1 had the highest number of precursors with 1,050 perfect matches across all chromosomes, Tae\_miR1436-2 has only 8, due to a single nucleotide difference that makes this sequence much less abundant in the genome (fig. 2A, last sequence in the alignment). In fact, Tae\_miR1436-1 alone accounts for ~20% (1,050 out of 5,490) of all precursors predicted for the 61 members of the miR1436 family (supplementary file S4, Supplementary Material online). Here, due to high redundancy of highly conserved MITE precursors, it is not possible to distinguish how many of these potential loci or which ones exactly are giving rise to Tae\_miR1436-1, thus illustrating one of the challenges inherent to the characterization of TE-born miRNAs (Li et al. 2011; Roberts et al. 2014; Qin et al. 2015). However, despite the very-high number of MITE loci encoding Tae\_miR1436-1, this sequence is never found in the uninfected sample, while it is always induced upon mildew infection. We therefore suggest that the miR1436 family, and particularly Tae\_miR1436-1, are highly relevant candidate miRNAs demonstrating the role of transposon domestication in regulating the wheat disease response.

MITEs are nonautonomous elements that are often found in UTRs, and are therefore frequently co-expressed with the nearby gene (Casacuberta et al. 1998; Santiago et al. 2002; Oki et al. 2008; Lu et al. 2012). Consistent with this observation, we found that miR1436 MITE copies, in particular those

encoding Tae\_miR1436-1, are enriched in gene transcripts (fig. 2C, supplementary note S2, Supplementary Material online, supplementary file S7, Supplementary Material online). Further analysis revealed that one potential miR1436 MITE copy is found in intron 18 of the Lr34 mildew resistance gene (Krattinger et al. 2009) (fig. 2D, GenBank: FJ436983.1), and encodes a highly conserved hairpin flanked by typical TIRs encoding the Tae\_miR1436-1 sequence (position 19,103, fig. 2E). We found another conserved copy in the 5′ UTR and same transcriptional orientation of a nucleotide-binding domain leucine-rich repeats (NLR) gene (GenBank: AYG86980.1), encoded in the NLR cluster of the powdery mildew resistance locus QPm.tut-4A (Jakobson et al. 2012; Janáková et al. 2019) (fig. 2F, G, GenBank: MG672525.1). This locus was introgressed from *Triticum militinae*, and together with Lr34, these results indicate that potential Tae\_miR1436-1 MITE precursors can be encoded in important resistance loci in several *tritici* lineages. Moreover, the Tae\_miR1436-1 MITE precursor is the only conserved “island” in an otherwise sequence-unrelated resistance locus (fig. 2H). We found no targets of Tae\_miR1436-1 within the Lr34 or the QPm.tut-4A loci, which further suggests this miRNA is contributing to the general response of wheat to mildew infection. We conclude that genome-wide distribution of thousands of MITEs encoding a conserved Tae\_miR1436-1 miRNA stem-loop, combined with co-occurrence within major mildew resistance loci in wheat and other close relatives, suggest that domestication of the Tae\_miR1436-1 MITE is an important contribution to the biotic stress-response in the *tritici* tribe.

To uncover the regulatory network of Tae\_miR1436-1, we performed a curated search for miRNA targets in the wheat genome including all 18 miRNA sequences commonly found in all infected samples (supplementary note S4, Supplementary Material online, supplementary tables S8–S10, Supplementary Material online, supplementary file S8, Supplementary Material online). We found that several genes were targeted by two or more miRNAs, that did not necessarily have similar sequences, and these could be classified in three groups (fig. 3A–C), with distinctly different functional profiles (fig. 3D–F, supplementary tables S8–S10, Supplementary Material online). Although Group\_2 targets consisted mainly of pentatricopeptide repeat (PPR) proteins, commonly involved in gene regulation (Manna 2015; supplementary table S9, Supplementary Material online), Group\_3 targets were functionally more diverse, but interestingly consisted mainly of genes encoding possible miR1127 precursors (supplementary table S10, Supplementary Material online). Group\_1 stands out as the largest network of miRNAs with shared targets, mainly overlapping with the Tae\_miR1436-1 predicted regulatory network (supplementary tables S8, S11, Supplementary Material online). These data suggest that mildew responsive miRNAs, with sometime distinctly different sequences, can form a coordinated miRNA response. Here, transposon-derived miRNAs seem to have the most functionally diverse regulatory networks, which can be exemplified by the 54 targets of Tae\_miR1436-1 (supplementary table S11, Supplementary Material online). Of these, one target



**Fig. 2.** Structural and genomic features of domesticated MITE precursors of wheat miRNAs. (A) Multiple sequence alignment of MITE elements showing conservation of the terminal inverted repeat sequences. (B) Stem-loop structures resulting in the formation of a perfect (upper sequence), or imperfect (lower sequence) duplex encoded by two members of the *DTT\_Pancho* MITE family. (C) Number of *Tae\_miR1436-1* precursors (darker color tone) from the *miR1436* family precursors (lighter color tone) encoded in intergenic sequences (upper panel) or within genes (lower panel). (D) Relative position of the conserved MITE *Tae\_miR1436-1* copy encoded in the *Lr34* locus, within intron 18 of the *Lr34* mildew resistance

(TraesCS3D01G141100) is encoding a putative small cysteine-rich Metallothionein 3-like (*TaeMt3*) protein (fig 3G, supplementary fig S5, Supplementary Material online), a group of proteins participating in an array of protective stress responses, with some corresponding to scavengers of reactive oxygen species (ROS) (Ruttkay-Nedecky et al. 2013). Prominent examples from agricultural crops are the GhMT3 from cotton (Xue et al. 2008), and OsMT1a and OsMT2b Metallothioneins from rice (Wong et al. 2004; Yamauchi et al. 2017). We then hypothesized that possible downregulation of *TaeMt3*, could result in derepression of the wheat oxidative burst response associated with cell-death signaling (Torres et al. 2006). We assessed the physiological effect of *Tae\_miR1436-1* on the expression of the *TaeMt3* target by comparing the expression of the miRNA precursor/target pair at 2 dpi and 7 dpi. Interestingly, we found fewer but transcriptionally much more active loci encoding miR1436-1 at 7 dpi (fig. 3H, I). Furthermore, 8 of the 17 miR1436 family loci upregulated at 7 dpi encode for the miR1436-1 variant (supplementary fig. 6, Supplementary Material online). Most importantly, we found significant downregulation of the *TaeMt3* target at 7 days compared with 2 days (fig. 3J). These data substantiate our hypothesis that selective transcriptional activation of miR1436-1 is physiologically meaningful, and that *Tae\_miR1436-1* and *TaeMt3* are a functional regulatory pair contributing to the wheat response to increasing mildew pressure. In order to further study the relevance of this regulatory pair to wheat immunity, we developed a series of functional assays to assess the suggested role of the *Tae\_miR1436-1/TaeMt3* pair in regulating cell-death (supplementary note S5, Supplementary Material online, supplementary figs. S7–S9, Supplementary Material online). Of these experiments, we established an assay where either *TaeMt3* or the empty vector control (pIPKb004) is expressed on two areas of the same *Nicotiana benthamiana* leaf. Then, cell death is induced in that same tissue using high density apoplastic infiltration with DC3000, a virulent strain from the *Pseudomonas syringae* pv. *tomato* pathogen (see Materials and Methods, supplementary note S5, Supplementary Material online). Results from three independent assays showed quantitative, partial suppression of the cell-death response on several leaves of the heterologous *N. benthamiana* host when *TaeMt3* was present in the mix (fig. 3K–M, supplementary fig. S9, Supplementary Material online). We therefore suggest that *Tae\_miR1436-1*-targeting of *TaeMt3* could possibly contribute to the derepression of cell death that is a particularly efficient immune response to the biotrophic mildews, which rely on living host tissues to survive.

In summary, we propose that 1) genome-wide distribution of thousands of MITE copies (Wicker et al. 2018), 2) frequent association of these with genes (Casacuberta et al. 1998; Santiago et al. 2002; Oki et al. 2008; Lu et al. 2012), combined with a functionally diversified network of MITE-born miRNA targets, might provide a selective advantage for frequent domestication of MITEs into miRNA genes relevant to biotic stress adaptation. We conclude that MITE domestication into miRNA precursors is probably an important evolutionary force driving miRNA innovation in the wheat small RNA response to invading pathogens, thus further substantiating the model proposed by Roberts and colleagues, arguing that miRNAs are initially formed from TEs (Roberts et al. 2014).

## Materials and Methods

### Biological Material, RNA Extraction and Sequencing

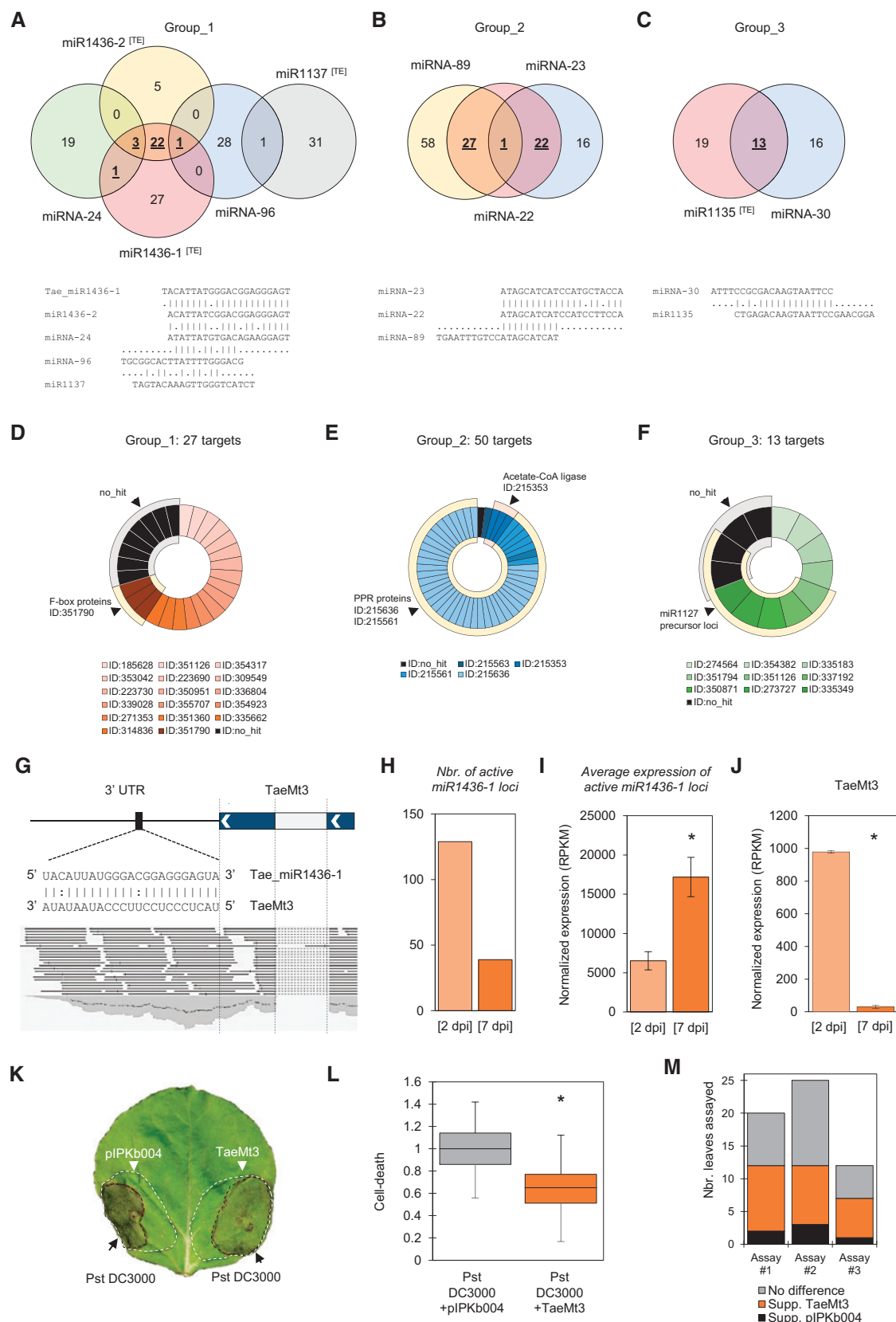
The *T. aestivum* cultivar “Chinese Spring” and the *B.g. tritici* isolates Bgt\_96224, Bgt\_94202, and Bgt\_JIW2 were used in this study. Plant growth and infection conditions were previously described by Praz et al. (2018). Briefly, ten days old detached leaf segments were infected with fresh spores from pure cultures of the virulent powdery mildew isolates Bgt\_96224, Bgt\_94202, and Bgt\_JIW2. The infected leaf segments were kept on benzimidazole agar plates at 20 °C and 70% humidity with a 16 h light/8 h dark cycle. For the microscopy image, wheat leaves were transiently transformed with a  $\beta$ -glucuronidase (GUS) reporter gene as previously described, whereas epiphytic mildew fungal structures were stained with coomassie blue as previously described in Nowara et al. (2010).

Small RNA samples were extracted after 2 days from leaf material infected with three different powdery mildew isolates and from uninfected leaves, whereas mRNA samples were extracted after 7 days only from leaves infected with the isolate Bgt\_96224. All RNA extractions were done using the miRNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. The quality of the extracted RNA was checked by gel electrophoresis and the purity was further assessed with the NanoDrop Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) based on the 260:280 ratios. For the small RNA samples, single-end-50 bp read libraries were generated with the TruSeq Small RNA Library Preparation Kit (Illumina) and sequenced with the Illumina HiSeq 2500 Sequencing System. For the mRNA samples, paired-end-150 bp read libraries were generated with the TruSeq Stranded mRNA library Preparation Kit (Illumina) and sequenced with the NovaSeq 6000 Sequencing System. The sequencing was carried out at the Functional Genomics

FIG. 2. Continued

gene. (E) Multiple sequence alignment of the Lr34 encoded MITE and other *Tae\_miR1436-1* MITE precursor sequences. Strong conservation of the stem-loop region delineated by the miRNA and miRNA start sequences (position 19 to 103) is illustrated. (F) Relative position of the conserved MITE *Tae\_miR1436-1* copy encoded in the NLR cluster of the QPm.tut-4A powdery mildew resistance locus from *Triticum militinae*. Relative position to the CC-NBS-LRR encoding gene is indicated. (G) Strong conservation of the stem-loop region between the QPm.tut-4A, the Lr34 MITE, and one additional *Tae\_miR1436-1* precursor is here illustrated. (H) Dotplot depiction showing that the putative MITE precursor of *Tae\_miR1436-1* is the only conserved sequence over large segments of the QPm.tut-4A and the Lr34 loci.





**FIG. 3.** Characterization of the endogenous wheat targets of mildew responsive miRNAs. (A–C) Venn diagrams illustrating the overlap between the targets of several miRNAs, possibly representing distinctly different coordinated responses. The miRNA forming group 1 (A), group 2 (B), and group 3 (C) are indicated (upper panel), and alignments of those same miRNAs are depicted (lower panels). (D–F) In silico prediction of functional protein domains using BLAST search against the NCBI conserved domain database (CDD). The identified protein superfamilies for group 1 (D), group 2 (E), and group 3 (F) are indicated by their PSSM-ID. (G) Description of the binding site of Tae\_miR1436-1 in the 3' UTR of the TaeMt3 target



Center Zürich (<http://www.fgcz.ch/>). Sequencing reads are available under the NCBI BioProject PRJNA553193 and PRJNA577532.

### Filtering of Raw Reads

Raw reads from the four small RNA libraries were cleaned and filtered based on sequencing quality, read length, and homology with non-coding RNAs (ncRNAs). The 3-prime adapter sequence was found using the “find\_3p\_adapter.pl” script (<http://sites.psu.edu/axtell/software/misc-tools/>; last accessed January 2017) with the sequence of a conserved and abundant plant miRNA sequence (ath-mir156a) with “-m TGACAGAAGAGAGTGAGCAC” option. Afterward, the raw small RNA sequencing reads were trimmed with FASTX Toolkit 0.0.13 ([http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html); last accessed January 2017) using the command “fastx\_clipper -l 18 -v -c -a TGGGAATTC -Q33.” The trimmed reads were filtered via “Seq\_filter.pl” ([https://github.com/dmb107/UTR.scripts/blob/master/Seq\\_filter.pl](https://github.com/dmb107/UTR.scripts/blob/master/Seq_filter.pl); last accessed January 2017) and only those with a length between 18 and 24 bp were kept for further downstream analysis. The reads were mapped with bowtie 1.0.0 with the option “-v 1” (Langmead et al. 2009) to the wheat ncRNA database ([ftp://ftp.ensemblgenomes.org/pub/plants/release-39/fasta/triticum\\_aestivum/ncrna/](ftp://ftp.ensemblgenomes.org/pub/plants/release-39/fasta/triticum_aestivum/ncrna/); last accessed January 2017) and the ones that aligned to snRNA, snoRNA, tRNA, and rRNA with zero or one mismatches were removed.

### Prediction of miRNAs

De novo annotation of miRNA loci was performed using ShortStack 3.6 with the recommended parameters for plants (Axtell 2013). The mapping software bowtie cannot handle large genomes, therefore the wheat genome (~17 Gb) was divided in 22 segments (each one corresponding to a chromosome) and indexed using “bowtie-build.” For the mapping of the clean small RNA-Seq reads on the chromosomes, bowtie 1.0.0 was used with the settings “-v 1 -m 50 -a -best -strata.” This procedure was followed for the libraries of the three infected samples and of the uninfected control. Homology-based identification of conserved miRNAs was performed as described by Alptekin et al. (2017). First, the sequences of experimentally validated miRNAs (miRBase, Release 21) were aligned to each of the cleaned small-RNA libraries with the perl script “SUMirFind\_smRNA.pl,” allowing up to two mismatches. In a second step, the small RNA reads

with homology to known miRNAs, were aligned back to the wheat genome with “SUMirFind.pl” without allowing any mismatches. “SUMirFold” was then used to test the presence of a potential hairpin shaped secondary structure. Finally, the genuineness of the putative miRNA loci was evaluated with “SUMirPredictor.” To determine which miRNAs are more likely to be expressed in response to pathogen infection, we selected the miRNAs that were only predicted upon infection (not present in the uninfected control) and that were commonly induced by three fungal isolates (94202, 96224, and JIW2).

### miRNA Precursor Loci Identification and Expression Studies

“SUMirLocator” was used to identify the genomic location of the premiRNAs that were previously predicted with “SUMirPredictor,” whereas information about the distribution of the “ShortStack” miRNA loci were already available from the results output of the pipeline. TE associated miRNAs were then identified by aligning the intergenic precursor sequences against the TEs (nrTREP17) database (botinst.uzh.ch/en/research/genetics/thomasWicker/trep-db.html; last accessed June 2018) using BLASTN with the following parameters: -evalue 1e-10 -qcov\_hsp\_perc 30. Additional information on the analysis of miRNA precursors are described in [supplementary note S2, Supplementary Material](#) online. In order to estimate the number of loci possibly located in genes transcripts, we generated a database consisting of the genomic sequences of all high-confidence wheat genes, defined as the sequence from gene start to gene stop, including exons, introns, and UTRs when available. The miR1127, miR1137, and miR1436 precursor sequences were aligned onto this database via BLASTN with the parameters: -min eval E-60, -max number of hits 1, -max number of alignments 100000, -gapped\_search disabled. The resulting blast results were filtered for hits with min of 98% identity, and manually curated for query coverage.

Wheat leaf segments from the reference wheat cultivar “Chinese Spring” were infected with the wheat powdery mildew reference isolate Bgt\_96224 as previously described (Praz et al. 2018). Samples were harvested 7 dpi, and applied to Illumina NovaSeq 150 pb paired end sequencing. Expression levels of miRNA precursors encoding for the MITE born families miR1127, miR1137, and miR1436, as well as specific precursors encoding Tae\_miR1436-1 were determined with

### Fig. 3. Continued

(two last exons are depicted). RNA sequencing reads demonstrating the existence of the predicted 3'UTR are depicted in the lower panel. (H) Number of expressed Tae\_miR1436-1 precursors at 2 and 7 dpi. (I) Average expression of active miR1436-1 loci at 2 and 7 dpi (see Materials and Methods, [supplementary note S2, Supplementary Material](#) online). Error bars indicate the standard error of the mean from three independent replicates (RPKM: reads per kilobase per million). (J) Downregulation of TaeMt3 at 7 dpi coinciding with higher induction of miR1436-1 precursors. Error bars indicate the standard error of the mean from three independent replicates. Statistical significance was assessed using the Student's *t*-test for paired data. *P* value < 0.05 is indicated by a star. (K) Restriction of the necrotic cell death in presence of TaeMt3 at 48 h post DC3000 treatment of *Nicotiana benthamiana* leaves (see [supplementary note S5, Supplementary Material](#) online for assay description). (L) Quantification of cell death intensity from the Pst DC3000 infiltrated leaf areas and expressing either the pIPKb004 empty expression vector, or TaeMt3. Cell death was scored in one assay from 15 independent leaf replicates (see Materials and Methods). (M) Stacked bar blot summary of the TaeMt3 cell death suppression phenotype observed from three independent assays, each consisting of 12–25 leaf replicates.

SalmonTE using standard parameters, which is specifically designed for the quantification of TEs in RNA-Seq data (github.com/LiuzLab/SalmonTE; last accessed September 2019). We determined expression levels in 1) RNA-Seq data which were generated at 2 dpi from the wheat cultivar “Chinese Spring” infected with the mildew isolate Bgt\_96224 (Praz et al. 2018), corresponding to the exact same genetic material used in our small RNA-Seq study, and 2) RNA-Seq libraries from three independent biological replicates from “Chinese Spring” infected with Bgt\_96224, but harvested at 7 dpi. Precursor expression studies are described in [supplementary note S2](#), [Supplementary Material](#) online.

### Target Prediction and Cloning

Complementary information on target prediction is provided in [supplementary note S4](#), [Supplementary Material](#) online. Briefly, we performed two independent analyses using psRNATarget (Dai et al. 2018), a plant specific small RNA target prediction software, with standard parameters, and the more generalist eukaryotic small RNA aligner GSTAR (v1.0; <https://github.com/MikeAxtell/GSTAR>; last accessed May 2018) with standard parameters against the wheat cDNA annotation (IWGSC RefSeq v1.0). The most relevant targets were selected based on three main parameters (MFE ratio: 0.73–1, Allen score: 0–4, and Expectation value: 0–3) and on the overlap between the two software. Targets of the subset of 18 mildew responsive miRNA are provided in [supplementary file S8](#), [Supplementary Material](#) online. TraesCS3D01G141100 was cloned from the genomic DNA for the wheat cultivar “Chinese Spring” using TOPO directional cloning according to the manufacturer (pENTR/D-TOPO, Thermofisher). The construct was verified using in-house Sanger sequencing.

### Suppression of *P. syringae* DC3000 Induced Cell Death

The TaeMt3 TOPO cloning cassette was recombined into the *Agrobacterium tumefaciens* expression vector pIPKb004 (Himmelbach et al. 2007) using Gateway LR clonase II (Invitrogen) according to the manufacturer and subsequently transformed into *A. tumefaciens* strain GV3101 using electroporation (1.44 kV, 25  $\mu$ F, 200  $\Omega$ ) (Bourras et al. 2015). *A. tumefaciens* mediated transient expression of TaeMt3, or the empty pIPKb004 vector in *N. benthamiana* leaves was achieved as previously described by Bourras et al. (2019). TaeMt3 was transiently expressed on one side of the leaf together with the tombusvirus P19 RNA silencing suppressor protein at a 4:1 ratio (Saxena et al. 2011). As a control, the empty pIPKb004 expression vector and P19 were expressed at the same ratio on the other side of the leaf. After 24 h, Pst-DC3000 cultures were infiltrated at OD<sub>600</sub> = 1.0 inside the area where TaeMt3 or the empty vector have been transiently expressed. Leaves were cut off the plants after another 24 h, and early cell-death was quantified using HSR imaging as previously described (Bourras et al. 2019). The leaves were incubated in the dark on water-soaked absorbent paper, in 100% humidity, 28 °C, in the dark, in an oven. After another 24–48 h, leaves were removed and scored for necrotic cell death development.

Pst-DC3000 cultures were prepared mostly following the same protocol as for *A. tumefaciens* with the following minor modification. Briefly, DC3000 cultures were grown in Luria broth (LB) medium supplemented with appropriate antibiotics overnight at 28 °C with 200 rpm shaking. The day after, cultures were harvested by centrifugation 3,300  $\times$  g, 5 min, resuspended in fresh LB medium without antibiotics and further incubated for 30 min at 28 °C with 200 rpm shaking. Bacteria were washed two times in an AS<sup>(-)</sup> medium (10 mM MES-KOH, pH5.6; 10 mM MgCl<sub>2</sub>; 200), which is a classical AS medium where acetosyringone is omitted. Finally, the cultures were adjusted in AS<sup>(-)</sup> medium to an OD of 1.0 before apoplastic infiltration of *N. benthamiana* leaves.

### Supplementary Material

[Supplementary data](#) are available at *Molecular Biology and Evolution* online.

### Acknowledgments

We would like to acknowledge Prof. Beat Keller (University of Zurich, Switzerland) and Prof. Bruce McDonald (ETH, Zurich) for their support. We would also like to acknowledge the technical support from Gerhard Herren, Helen Zbinden, Karl Huwiler from the University of Zurich, and Dr Sirisha Aluri from the Functional Genomics Center Zurich (FGCZ). We also thank the Swiss National Science Foundation (grant number 31003A\_163325) for the financial support, and the University of Zurich Research Priority Program (URPP).

### Authors Contribution

M.P., T.W., and S.B. wrote the paper. M.P., T.W., and S.B. designed the research. M.P., C.R.P., A.S.V., and S.B. designed lab experiments. M.P., C.R.P., L.K., C.K., L.K.S., M.S., V.W., A.S.V., and S.B. did lab experiments. M.P., C.R.P., T.W., and S.B. did bioinformatics analyzes. M.P., T.W., and S.B. coordinated the research.

### References

- Alptekin B, Akpinar BA, Budak H. 2017. A comprehensive prescription for plant miRNA identification. *Front Plant Sci.* 7:2058.
- Axtell MJ. 2013. ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA* 19(6):740–751.
- Bourras S, Kunz L, Xue M, Praz CR, Müller MC, Kälin C, Schläfli M, Ackermann P, Flückiger S, Parlange F, et al. 2019. The AvrPm3-Pm3 effector-NLR interactions control both race-specific resistance and host-specificity of cereal mildews on wheat. *Nat Commun.* 10(1):2292.
- Bourras S, McNally KE, Ben-David R, Parlange F, Roffler S, Praz CR, Oberhaensli S, Menardo F, Stirnweis D, Frenkel Z, et al. 2015. Multiple avirulence loci and allele-specific effector recognition control the Pm3 race-specific resistance of wheat to powdery mildew. *Plant Cell.* 27(10):2991–3012.
- Budak H, Kantar M, Bulut R, Akpinar BA. 2015. Stress responsive miRNAs and isomiRs in cereals. *Plant Sci.* 235:1–13.
- Bureau TE, Wessler SR. 1994. Stowaway: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. *Plant Cell.* 6:907–916.
- Cagliani R, Sironi M. 2013. Pathogen-driven selection in the human genome. *Int J Evol Biol.* 2013:1.

- Casacuberta E, Casacuberta JM, Puigdomènech P, Monfort A. 1998. Presence of miniature inverted-repeat transposable elements (MITEs) in the genome of *Arabidopsis thaliana*: characterisation of the Emigrant family of elements. *Plant J.* 16(1):79–85.
- Dai X, Zhuang Z, Zhao PX. 2018. psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* 46:49–54.
- Djami-Tchatchou AT, Sanan-Mishra N, Ntushelo K, Dubery IA. 2017. Functional roles of microRNAs in agronomically important plants—potential as targets for crop improvement and protection. *Front Plant Sci.* 8:378.
- Himmelbach A, Zierold U, Hensel G, Riechen J, Douchkov D, Schweizer P, Kumlehn J. 2007. A set of modular binary vectors for transformation of cereals. *Plant Physiol.* 145(4):1192–1200.
- Jakobson I, Reis D, Tiidema A, Peusha H, Timofejeva L, Valárik M, Kladičková M, Šimková H, Doležel J, Järve K. 2012. Fine mapping, phenotypic characterization and validation of non-race-specific resistance to powdery mildew in a wheat-*Triticum militinae* introgression line. *Theor Appl Genet.* 125(3):609–623.
- Janáková E, Jakobson I, Peusha H, Abrouk M, Škopová M, Šimková H, Šafář J, Vrána J, Doležel J, Järve K, et al. 2019. Divergence between bread wheat and *Triticum militinae* in the powdery mildew resistance QPm.tut-4A locus and its implications for cloning of the resistance gene. *Theor Appl Genet.* 132(4):1061–1072.
- Jung JH, Park CM. 2007. MIR166/165 genes exhibit dynamic expression patterns in regulating shoot apical meristem and floral development in *Arabidopsis*. *Planta* 225(6):1327–1338.
- Karasov TL, Horton MW, Bergelson J. 2014. Genomic variability as a driver of plant-pathogen coevolution? *Curr Opin Plant Biol.* 18:24–30.
- Krattinger SG, Lagudah ES, Spielmeier W, Singh RP, Huerta-Espino J, McFadden H, Bossolini E, Selter LL, Keller B. 2009. A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323(5919):1360–1363.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10(3):R25.
- Li Y, Li C, Xia J, Jin Y. 2011. Domestication of transposable elements into microRNA genes in plants. *PLoS ONE* 6(5):e19212.
- Lu C, Chen J, Zhang Y, Hu Q, Su W, Kuang H. 2012. Miniature inverted-repeat transposable elements (MITEs) have been accumulated through amplification bursts and play important roles in gene expression and species diversity in *Oryza sativa*. *Mol Biol Evol.* 29(3):1005–1017.
- Manna S. 2015. An overview of pentatricopeptide repeat proteins and their applications. *Biochimie* 113:93–99.
- Mariño-Ramírez L, Lewis KC, Landsman D, Jordan IK. 2005. Transposable elements donate lineage-specific regulatory sequences to host genomes. *Cytogenet Genome Res.* 110(1–4):333–341.
- Nowara D, Gay A, Lacomme C, Shaw J, Ridout C, Douchkov D, Hensel G, Kumlehn J, Schweizer P. 2010. HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen *Blumeria graminis*. *Plant Cell* 22(9):3130–3141.
- Oki N, Yano K, Okumoto Y, Tsukiyama T, Teraishi M, Tanisaka T. 2008. A genome-wide view of miniature inverted-repeat transposable elements (MITEs) in rice, *Oryza sativa* ssp. *japonica*. *Genes Genet Syst.* 83(4):321–329.
- Piriyapongsa J, Jordan IK. 2008. Dual coding of siRNAs and miRNAs by plant transposable elements. *RNA* 14(5):814–821.
- Piriyapongsa J, Mariño-Ramírez L, Jordan IK. 2007. Origin and evolution of human microRNAs from transposable elements. *Genetics* 176(2):1323–1337.
- Praz CR, Menardo F, Robinson MD, Müller MC, Wicker T, Bourras S, Keller B. 2018. Non-parent of origin expression of numerous effector genes indicates a role of gene regulation in host adaption of the hybrid triticale powdery mildew pathogen. *Front Plant Sci.* 9:49.
- Qin S, Jin P, Zhou X, Chen L, Ma F. 2015. The role of transposable elements in the origin and evolution of microRNAs in human. *PLoS ONE* 10(6):e0131365.
- Ravichandran S, Ragupathy R, Edwards T, Domaratzi M, Cloutier S. 2019. MicroRNA-guided regulation of heat stress response in wheat. *BMC Genomics* 20(1):488.
- Rebollo R, Romanish MT, Mager DL. 2012. Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu Rev Genet.* 46(1):21–42.
- Roberts JT, Cardin SE, Borchert GM. 2014. Burgeoning evidence indicates that microRNAs were initially formed from transposable element sequences. *Mob Genet Elements.* 4(3):e29255.
- Ruttkey-Nedecky B, Nejdil L, Gumulec J, Zitka O, Masarik M, Eckschlagner T, Stiborova M, Adam V, Kizek R. 2013. The role of metallothionein in oxidative stress. *IJMS* 14(3):6044–6066.
- Santiago N, Herráiz C, Ramón Goñi J, Messegueur X, Casacuberta JM. 2002. Genome-wide analysis of the Emigrant family of MITEs of *Arabidopsis thaliana*. *Mol Biol Evol.* 19(12):2285–2293.
- Saxena P, Hsieh YC, Alvarado VY, Sainsbury F, Saunders K, Lomonosoff GP, Scholthof HB. 2011. Improved foreign gene expression in plants using a virus-encoded suppressor of RNA silencing modified to be developmentally harmless. *Plant Biotechnol J.* 9(6):703–712.
- Sorin C, Declerck M, Christ A, Blein T, Ma L, Lelandais-Brière C, Njo MF, Beeckman T, Crespi M, Hartmann C. 2014. A miR169 isoform regulates specific NF-YA targets and root architecture in *Arabidopsis*. *New Phytol.* 202(4):1197–1211.
- Torres MA, Jones JDG, Dangl JL. 2006. Reactive oxygen species signaling in response to pathogens: figure 1. *Plant Physiol.* 141(2):373–378.
- Vicient CM, Casacuberta JM. 2017. Impact of transposable elements on polyploid plant genomes. *Ann Bot.* 120(2):195–207.
- Wicker T, Gundlach H, Spannagl M, Uauy C, Borrill P, Ramírez-González RH, De Oliveira R, Mayer KFX, Paux E, Choulet F. 2018. Impact of transposable elements on genome structure and evolution in bread wheat. *Genome Biol.* 19(1):103.
- Wong HL, Sakamoto T, Kawasaki T, Umemura K, Shimamoto K. 2004. Down-regulation of metallothionein, a reactive oxygen scavenger, by the small GTPase OsRac1 in rice. *Plant Physiol.* 135(3):1447–1456.
- Wu G, Park MY, Conway SR, Wang JW, Weigel D, Poethig RS. 2009. The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell* 138(4):750–759.
- Xu MY, Zhang L, Li WW, Hu XL, Wang MB, Fan YL, Zhang CY, Wang L. 2014. Stress-induced early flowering is mediated by miR169 in *Arabidopsis thaliana*. *J Exp Bot.* 65(1):89–101.
- Xue T, Li X, Zhu W, Wu C, Yang G, Zheng C. 2008. Cotton metallothionein GhMT3a, a reactive oxygen species scavenger, increased tolerance against abiotic stress in transgenic tobacco and yeast. *J Exp Bot.* 60(1):339–349.
- Yamauchi T, Fukazawa A, Nakazono M. 2017. METALLOTHIONEIN genes encoding ROS scavenging enzymes are down-regulated in the root cortex during inducible aerenchyma formation in rice. *Plant Signal Behav.* 12(11):e1388976.
- Yan Y, Zhang Y, Yang K, Sun Z, Fu Y, Chen X, Fang R. 2011. Small RNAs from MITE-derived stem-loop precursors regulate abscisic acid signaling and abiotic stress responses in rice. *Plant J.* 65(5):820–828.